DOCUMENT RESUME

ED 143 199 FL 008 296

AUTHOR Killingley, Siew-Yue
TITLE Internal Structure of the Cantonese Word and General Problems of Word Analysis in Chinese. York Papers in Linguistics, No. 7.
INSTITUTION York Univ. (England). Dept. of Language.
PUB DATE 75
NOTE 31p.
AVAILABLE FROM York Papers in Linguistics, Department of Language, University of York, Heslington, York, 4015DD, England

EDRS PRICE MF-$0.83 HC-$2.06 Plus Postage.
DESCRIPTORS *Cantonese; Chinese; Language Research; Language Usage; Literacy; *Mandarin Chinese; *Morphology (Languages); Sino Tibetan Languages; *Sociolinguistics; *Structural Analysis; Written Language

ABSTRACT

This paper: (1) examines the internal word structure of Cantonese in terms of root, stem and affix; (2) discusses general problems of word analysis in Chinese with reference to Cantonese and Mandarin; (3) examines the issue of distorted translations from Mandarin into English resulting from faulty word analysis; and (4) looks briefly at the problem of distinguishing between word and phrase in Cantonese. The topics are discussed in the light of the sociolinguistic situation in Singapore and Malaya, in which it is not unusual to find a native speaker of Cantonese or some other Chinese language who is literate only in Mandarin. The latter serves as a lingua franca for educated speakers in speech and writing. The descriptive model used in this paper is a modified tagmemic one. (CLK)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Siew-Yue Killingley
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND THE ERIC SYSTEM CONTRACTORS."

ED143199

INTERNAL STRUCTURE OF THE CANTONESE WORD AND GENERAL PROBLEMS OF WORD ANALYSIS IN CHINESE*

Siew-Yue Killingley

(St. Mary's College of the Sacred Heart,
Newcastle upon Tyne)

1.0. Introduction

1.1. Complexities of the Chinese sociolinguistic situation in Malaya and Singapore. In this paper, I hope to examine the internal word structure of Cantonese in terms of root, stem and affix, to discuss general problems of word analysis in Chinese with reference to Cantonese and Mandarin, to examine the problem of distorted translations from Mandarin into English arising from faulty word analysis, and to look briefly at the problem of distinguishing between word and phrase in Cantonese.

The inclusion of Cantonese and Mandarin in the same paper may appear unwarrantably confusing to anybody who is not familiar with the Chinese sociolinguistic scene in Singapore and Malaya, in which innumerable educated Chinese speakers are orally bilingual (in Mandarin and one other Chinese language), but in writing, mainly monolingual (in Mandarin). The reasons behind such a situation are educational and cultural; Chinese schools use Mandarin as the language of instruction and Mandarin is regarded as a fitting language for the arts. In such a sociolinguistic setting, as far as Malayan and Singaporean Chinese are concerned, it is not unusual to find a native speaker of Cantonese, Hakka, Hokkien, Tiechiu, Hainanese, Kwongsai, Hockchui, Henghwa or Hokchia, who is only literate in Mandarin, the last serving as a lingua franca for an educated body of speakers in speech and used in writing.

The sociolinguistic situation mentioned above is not truly comparable to one in England in which a child speaks a non-standard dialect of English at home while speaking (and writing) Standard English at school. Although certain parallels obtain (such as choice of language depending on whether the child is at school or at home), a very important difference is that each of the non-Mandarin languages is fully acceptable to the Chinese as formal languages for oral purposes, but habit often governs the choice of Mandarin as a written language.

There is yet a further complication. Any non-Mandarin Chinese language can be used as a medium of instruction for reading Classical and other texts, and this is indeed done in the case of schools where Cantonese, for example, is used as the medium of instruction, or in some cases of private tuition at home when parents might prefer teaching to be done in a non-Mandarin language (though this is somewhat unusual). In such cases, although it is perfectly possible for a person to write Cantonese, we are by no means certain that users of Chinese think of the same thing when referring to 'written Cantonese'. In traditional practice, a piece of Mandarin can be read off, the reader using a special kind of Cantonese 'literary' pronunciation and changing the vocabulary and syntax of the read text (while reading it)

FL008296

I	II	III	IV	V	VI
Spoken Mandarin	*Written Mandarin & traditional written Cantonese*	*Pronunciation of Cantonese based on Mandarin*	*Spoken Colloquial Cantonese*	*'Missionary-favoured' written Cantonese*	*Pronunciation of 'Missionary-favoured' written Cantonese*
1 /wɔ3mən2 bu2ʂI1/ 'we are not'	我 們 不 是	/ŋɔ3mun5 bʌd1si4/	/ŋɔ3dei4 m5hai4/	我 哋 唔 像	/ŋɔ3dei4m5hai4/
2 /ta1də fu1tsin1/ 'his father'	他 的 父 親	/køj3dig1 fu1tsʌn1/	/køj3gə lou3dau4/	(a) 佢 嘅 老 嗰 or (b) 佢 嘅 父 親	(a) /køj3gə lou3dau4/ or (b) /køj3gə fu1tsʌn1/
3 /tsy2tz/ 'orange' [the fruit]	橘 子	/tsaŋ2tzi2/	/tsaŋ2/	橙	/tsaŋ2/
4			/wu1li1ma5tsa5/[3] 'muddled'		

in order to accommodate it to Cantonese vocabulary and syntax.[1] Some people would regard this combination of written Mandarin and Cantonese-type pronunciation as 'written Cantonese' (see columns II and III in Table I).

Alternatively, 'written Cantonese' can refer to a non-Mandarin written form reflecting a great deal of spoken Cantonese (but still containing a certain literary element not found in the spoken language), a written form favoured by missionary scholars like O'Melia and Cowles (see column V in Table I). This written form, useful to Chinese scholarship in many ways, has not, however, gained the general approbation of every Chinese scholar, and does not, as yet, enjoy the total acceptance that written Mandarin enjoys. Moreover, many normal colloquial Cantonese forms are not represented in writing even in this kind of supposedly colloquial Cantonese (see Table I, no. 4). Table I illustrates the uncertain position of 'written Cantonese'.

We see, therefore, that in talking about any Chinese language, there is a less straightforward correspondence between the spoken and the written form in the case of a non-Mandarin language than in the case of Mandarin, and written Mandarin has a greater sense of reality to the Singaporean and Malayan Chinese. For these reasons, although the main emphasis of this paper is on Cantonese, in the discussion of the relationships between the character, the syllable, the morpheme and the word in Chinese (4.0.), when both the spoken and the written language have to be considered, I have chosen to illustrate my general points with a few additional examples from Mandarin (Table III), as well as from Cantonese (Table II). In that section, common problems concerning the identification of linguistic units can be seen in the two languages.

It is now becoming increasingly acceptable to regard Chinese as a polysyllabic language, although as recently as 1964, Robins (1964:333) referred to Chinese as a monosyllabic language, albeit not an extreme example of such a language. Lyons (1968:188) sums up the more recent view that 'nowadays scholars appear to agree that many Chinese words are composed of more than one morpheme.' This is borne out, for instance, in Chao (1968) and Dow (1972); we are beginning to view Chinese in a new light, for Lin (1972:xxi) goes so far as to state that 'eighty percent of words in the spoken tongue are bisyllabic' and that 'the overwhelming character of polysyllabism in modern Chinese must be apparent to everyone.'

However, the attention of the majority of scholars has been on Mandarin, so that in many cases we can safely substitute the word 'Mandarin' for 'Chinese' whenever Chinese is mentioned in the field of linguistics. There is no standard work on the Cantonese word as a grammatical unit, and no modern descriptive work has been done on the Cantonese word in Malaya and Singapore, so it is hoped that this paper will be of some particular as well as general interest to scholars in the field.

1.2. The descriptive model and the romanization systems used The descriptive model used in this paper is a modified tagmemic model. Although actual tagmemes are not set up, the levels of the grammatical hierarchy have been found useful in sorting out problems of word analysis which cannot be resolved by looking at internal structure alone.

For the Mandarin examples (except when they occur in non-Wade-Giles quotations), the Wade-Giles system of romanization found in Mathews (1931), is used. Whenever a quoted example occurs in some other system of romanization,

the Wade-Giles equivalent is supplied within square brackets. For the Cantonese examples (except in the case of quoted examples using some other system of romanization), the system of romanization used is one which I devised and used in Killingley (1972). When any other system of romanization is used in a quoted Cantonese example, that example is spelt in my own system of romanization and supplied within square brackets (as in fn.1). It is hoped that reducing the many different systems of romanization to two will make the examples more readable.

The two tone systems, represented by the figures 1-4 in the Mandarin and the figures 1-5 in the Cantonese, must be kept distinct from each other. In both these systems, the neutral tone is left unmarked. The two tone systems are tabulated below:

Mandarin		*Cantonese*	
Tone 1	level tone	Tone 1	high level tone
Tone 2	rising tone	Tone 2	mid-rising tone
Tone 3	low rising tone	Tone 3	mid-level tone
Tone 4	falling tone	Tone 4	low level tone
		Tone 5	very low level tone

2.0. *Definition of the word*

The Cantonese word can be defined according to both external and internal criteria. Externally, the word typically (but not always) fills a slot on the phrase level; it does not always fill a phrase level slot because of the possibility of level-skipping. For instance, an adjective word such as *hung5* 'red', typically fills a noun phrase slot as a modifier, e.g. in the noun phrase *hung5 haay5* 'red shoes'. However, it can skip the phrase and clause levels and occur on the highest grammatical level of sentence, given an appropriate final intonation contour. *hung5*, in such a case, would be an elliptical sentence, its meaning depending on the wider context of its occurrence.

Words like *hung5* are *syntactically independent*, i.e., they can occur alone (Bloomfield's minimum free form), and, with the appropriate intonation, they regularly level-skip to become base fillers of elliptical sentences.

However, other words in Cantonese are *syntactically dependent*, i.e., they typically occur as fillers of phrases and only very rarely level-skip in order to fill slots in elliptical sentences. When they do, it is almost as if they were occurring in the form of citations: '*yi* jaw4 *go2*?' '*yi1*.' ('*This* or *that*?' '*This*'.) Such words e.g., the deictic *yi1* 'this', the classifier *go3* and the numeral *gey2* 'a few', under normal circumstances, never occur alone, but only as fillers of the limiting phrase, as in *yi1 go3* 'this one' and *yi1 gey2 go3* 'these few'.

Further, the word is either a) capable of expansion or b) not capable of expansion at word level. Words like *hung5* 'red' are potentially expandable, e.g. *hung5hung5* 'reddish', *hung5go3* 'redder'; simple root words like *yi1*, *go3* and *gey2* are not potentially expandable at word level either by the addition of roots or through affixation. This definition of word, therefore,

differs from Longacre's in that it includes simple roots which are not capable of word-level expansion. Since the chief concern of this paper is the internal structure of words in Cantonese, words which are not capable of expansion are no longer alluded to in the main argument of this paper.

2.1. Root, stem and affix Roots and stems occupy *nuclear* positions in the word while affixes occupy *satellite* positions in the word. Roots are either *free* or *bound* morphemes. The *free root* can also be called *simple stem* and it is a near-equivalent to Bloomfield's *minimum free form* in that it can nearly always occur by itself. When the free root occurs by itself, it is coterminous with *word*, and skips a couple of levels in the grammatical hierarchy in order to fill a sentence-level slot.

As mentioned earlier, syntactically dependent words like the deictic *yi1*, although free roots, cannot normally occur by themselves, but only as citations. Normally, all free roots fill word-level slots and phrase-level slots.

Bound roots can neither occur by themselves nor can they fill phrase-level slots. Normally, they only fill word-level slots. Bound roots are not affixes in that they do not regularly mark word classes. A bound root may occur in several lexical items and have a similar meaning in each lexical item, but these lexical items may belong to different word classes. The criterion for regarding a bound root as a morpheme[4] (and not as a phonological syllable) is whether it occurs frequently as a bound form (with the same basic meaning) in compound words.

The stem is either *simple*, in which case it consists of one free root, or it is *compound*, in which case it consists of more than one root. *Compound stems* include reduplicated roots.

At one point, then, *free root*, *simple stem* and *word* are all coincident. This happens when a free root which is potentially expandable occurs by itself as a simple stem. For example, the free root (and word) *yan5* 'mankind; man; person', is also a simple stem because it is potentially expandable. It can, for instance, take the pronominal suffix *-dey4* (see Table II, no. 8a) to form the impersonal pronoun word *yan5dey4* 'oneself', and it can be reduplicated to form a new noun word of two free roots *yan5yan5* 'everybody'.

The bound root, however, is only part of a word. For example, the adjective word *hab4faad3* 'legal' is made up of a free verb root *hab4* 'agree' and a bound root *faad3*, which has the general meaning 'law'. The bound root *faad3* is not written with a hyphen before or after it to indicate its bondage because it can occur at the beginning of a word compound or at the end of a word compound. In *hab4faad3*, it occurs at the end of the compound, but in *faad3gun1* 'judge', it occurs at the beginning of the compound, *gun1* being a free root (and word) which means 'official, magistrate'.

Thus we see that not every root is a stem, and that the two terms *root* and *stem* are needed in order to distinguish between the two concepts. The affix is like a bound root in that it cannot occur alone, but, unlike

a bound root, it regularly marks word classes. It is also like a free root in that it does not only fill slots on the word level; it can fill slots on the phrase level[5]. It is not, however, treated as a root in that it is a 'recurring partial' (Longacre 1964:9.0.), giving repetitive information such as verbal aspect, relativity and negation. The affix may be a prefix, an infix, a supr fix or a suffix, the most common being the suffix and the least common being the suprafix. The prefix only occurs with some nouns (e.g., '*daay4*hog4sang1' 'undergraduate'), the ordinal numeral (e.g. '*daay4*yad1' 'first')[6] and the negative word (e.g., '*m5*how2' 'don't'). The infix only occurs in *conjunct verbs* (3.2) (e.g., 'saw4*haa3*siong1' 'to sustain wounds for a time') and in the interrogative phrase (e.g., 'ley5-*m5*-ley5' are (you etc.) coming?'). The suprafix takes the form of tone modification on one of the roots of a two-root construction - in nouns, adjectives, adverbs or temporal words.

3.0. Word types

There are altogether five word types, divided according to their internal structure. Each word type embraces many different word classes and therefore includes words with different syntactic functions e.g., in 3.1., an adjective word of one free root, *kwaay5* 'naughty', is grouped together with an adverb word of one free root, *juen1dang1* 'deliberately'. Similarly, in 3.2., a noun word of more than one free root, *saam1doy2* 'shirt pocket', is grouped together with a verb word of more than one free root, *saw4fu2* 'suffer'. The reason for grouping together words of different syntactic function in this way can be justified on the grounds that this paper is concerned with word-level *constructions* and not with syntactic distribution.

These five word types are listed and exemplified below; their word class membership is indicated in brackets in each case in order to give some idea of their syntactic function, although that is not a primary concern of this paper. The syntactic function of words in phrases and the class membership of words are dealt with in chapters 9 and 10 of Killingley (1972).

3.1. Monomorphemic words Monomorphemic words consist of one free root which may be either monosyllabic or polysyllabic. A full argument as to why these words are treated as monomorphemic words is given in 4.1. and in 5.0. (General problems of word analysis): in brief, they are treated as monomorphemic words firstly, for the same reasons that, synchronically, *butterfingers, daddylonglegs, bishop* and *sinecure* would be regarded as monomorphemic words by many linguists today[7]. In English, we are not likely (except when joking) to create words like **buttertoes* (on the analogy of *butterfingers*) and **mummylonglegs* (on the analogy of *daddylonglegs*). However, we could, conceivably, coin words like *superbutterfingers* and *she-daddylonglegs*, treating *butterfingers* and *daddylonglegs* as unanalysable wholes. In *bishop* and *sinecure*, as there are historically no English morphemes in their construction, it is easier to see that they ought to be considered monomorphemic synchronically. Secondly, the words below are treated as monomorphemic words for the same reason that English *parrot* would not be analysed into *par + rot* and English *restaurant* would not be analysed into *rest + (au) + rant*.

1	kwaay5	'naughty'
2	juen1dang1	'deliberately'
3	si1man5[8]	'elegant; well-mannered'
4	haam3baa3laang3	'all'
5	yi5gaa1	'now'
6	wu1li1maa5caa5	'muddled'

3.2. Polymorphemic words of more than one free root or stem The second type of word is made up of one free root (or stem) in combination with another free root (or stem) in a modifier-head endocentric construction. The individual meanings of the roots and stems are provided within brackets. Roots are spelt with small letters while stems are spelt with capital letters.

Among the word-types in this section are a sub-class of intransitive verbs which were mentioned in 2.1. These verbs can be called the *conjunct verb* and a few of these permit a downgraded limiting phrase to interrupt the two free roots which go to make up the verb. Example (6) below, *saw4siong1* 'to be wounded', is such a verb. This verb word can be interrupted by *yad1 did1* 'a little', a downgraded limiting phrase, and the result is a special case of word. This violates, to a certain extent, the criterion of uninterruptability, but this is an exception to the general rule of internal stability of the word. This phenomenon may be compared to the case of interruption within words by expletives in English, but it must be borne in mind that in Cantonese, within this small area of verbs, the limiting phrase *regularly* interrupts the intransitive verb. On the whole, however, the conjunct verb is not interruptable by a downgraded limiting phrase. Examples (7), (8) and (9) below are such verbs.

1	saam1doy2	'shirt pocket'	saam1 'clothes'	+	doy2 'pocket'
2	din4ce1	'car'	din4 'electricity	+	ce1 'vehicle, chariot'
3	hung5sab4ji4ce1	'ambulance'	HUNG5SAB4JI4 'Red Cross'	+	ce1 'vehicle, chariot'
4	taay3gig4kuen5taw5	'Chinese art of self-defence'	TAAY3GIG4 'Taoism'	+	KUEN5TAW5 'fist'
5	jing3ji4fo1hog4	'political science'	JING3JI4 'politics'	+	FO1HOG4 'science'
6	saw4siong1	'to be wounded'	saw4 'sustain'	+	siong1 'wound'
7	yam2sing3	'drink merrily'	yam2 'drink'	+	sing3 'surname'
8	saw4fu2	'suffer'	saw4 'sustain'	+	fu2 'hardship'
9	coed1saan1	'(the dead) go on (their) funeral procession	coed1 'go out'	+	saan1 'mountain'

3.3 Dimorphemic words of one bound root and one free root In some compounds of this kind, there may be a suprafix in the form of tone modification added to one of the roots.

It was said earlier (2.1.) that the criterion for regarding a bound root as a separate morpheme was whether it occurred elsewhere in a bound form with the same basic meaning. This criterion can be tested in the words of this section. The bound root in each case is underlined. In the first three examples, the bound root is *faad3*, which has the general meaning 'law'. Examples (4) and (5) involve tone modification on the second root, a change from tone 5 to tone 2 in each case.

1	hab4*faad3* 'legal'	hab4 'to cooperate'	+	faad3 'law'
2	faan4*faad3* 'to violate the law'	faan4 'to go against the established order'	+	faad3 'law'
3	*faad3*gun1 'a judge'	faad3 'law'	+	gun1 'high official'
4	din4*yaw2* 'petrol'	din4 'electricity'	+	yaw5 generic term for oil
5	to1*haay2* 'slipper'	to1 'to drag'	+	haay5 'shoe'

3.4. Words made up of simple/compound stem + segmental affix(es) This word-type is illustrated with examples from adjectives and verbs; segmental affixes are underlined in the words and are spelt with a hyphen preceding them. Compound stems include reduplications.

With the small area of conjunct verbs mentioned earlier, i.e., those which regularly permit interruption by a downgraded limiting phrase, one or more aspectual affixes occur as infixes. The downgraded limiting phrase, in such a case, occurs after the aspectual affix or affixes. This is exemplified in no. (8) below, the limiting phrase being enclosed within brackets.

In the rest of the uninterruptable conjunct verbs, one or more aspectual affixes occur as infixes (see examples (9) and (10) below). In the non-conjunct verbs, the aspectual affix occurs as a suffix, as demonstrated in examples (11) and (12) below.

1	hung5*did1* 'a little redder'	hung5 'red'	+	-did1 'a little more'
2	caw2gwaay3*did1* 'a little more ugly'	CAW2GWAAY3 'ugly'	+	-did1 'a little more'
3	hung5*go3* 'redder'	hung5 'red'	+	-go3 comparative suffix
4	caw2gwaay3*go3* 'uglier'	CAW2GWAAY3 'ugly'	+	-go3 comparative suffix
5	ley5*jo2* '(he) has come'	ley5 'come'	+	-jo2 perfective aspect

6	hung5hung5*dey2* 'rather red'	HUNG5HUNG5 'reddish'	+	-dey2 'a little'				
7	ceng1ceng1*dey2* 'greenish'	CENG1CENG1 'greenish'	+	-dey2 'a little'				
8	daa2*yuen5jo2*(gey2 nin5) jiong3 'to have completed a war (of a few years' duration)'	DAA2JIONG3 'to go to war'	+	-yuen5 completive aspect	+	-jo2 perfective aspect		
9	yam2*jo2*sing3 'to have drunk merrily'	YAM2SING3 'drink merrily'	+	-jo2 perfective aspect				
10	coed1*yuen5jo2*saan1 '(the dead) have finished (with the whole business of) going on their funeral procession'	COED1SAAN1 '(the dead) go on their funeral procession'	+	-yuen5 completive aspect	+	+	-jo2 perfective aspect	
11	sig4*yuen5jo2* 'to have finished eating'	sig4 'eat'	+	-yuen5 completive aspect	+	-jo2 perfective aspect		
12	hab4*faan1maay5jo2* 'to have finished closing again'	hab4 'close'	+	-faan1 reciprocal aspect	+	-maay5 completive aspect	+	-jo2 perfective aspect

3.5 Words resulting from reduplication Certain words, namely, subclassifications of the adjective word, temporal word, adverb word, classifier word and locative word, may be constructed by different kinds of reduplication. In certain cases, a suprafix may be added to one of the reduplicated roots.

Straightforward reduplications of free roots have been given as examples of compound stem to illustrate another word-type in 3.4., viz., *hung5hung5* 'reddish' and *ceng1ceng1* 'greenish'. Such reduplications, therefore, are not further exemplified here. However, there are other kinds of reduplication; for instance, in certain monomorphemic or dimorphemic adjectives of two syllables, reduplication of the syllables takes place; each syllable is reduplicated in turn to form a resultant *adverb* word of four syllables (see examples (1), (2) and (3) below). Unlike (1), which illustrates a case of one-root reduplication, the reduplicated forms in (2) and (3) exemplify a kind of word formation which is still productive, as attested by the existence of compounds like *siw2hey3* 'easily offended, petty', *say3saam1* 'careful', *haag3to3waa2* 'polite greetings' and *haag3cing5* 'hospitality'. Examples (4) and (5) are unusual in that they are formed by reduplicating the *bound* locative root.

	Reduplicated adverb word	*Disyllabic adjective word of one or two roots*
1	si1si1man5man5[9] 'elegantly; politely'	si1man5 'elegant; polite'
2	haag3haag3hey3hey3 'politely; with reticence'	haag3hey3 'polite; reticent'

3	siw2siw2sam1sam1 'carefully'	siw2sam1 'careful'
	Reduplicated locative word	*Words with the bound locative root*
4	sue3sue3 'everywhere'	go2sue3 'there' yi1sue3 'here' bin1sue3 'where'
5	dow4dow4 'everywhere'	go2dow4 'there' yi1dow4 'here' bin1dow4 'where'

4.0 Some relationships between the character, the syllable, the morpheme and the word in Chinese In Tables II and III, I have attempted to set out some of the possible relationships between the character, the syllable, the morpheme and the word in Cantonese and Mandarin respectively, relationships by no means clearly agreed upon by scholars of Chinese. Cowles, for instance, although fully aware of the existence of free and bound forms, is not consistent in recording them as such (see remarks below on nos. 6c and 6d, 8a and 8b, and 13a and 13b in Table II); Chao, although a pioneer in establishing polysyllabism and polymorphism in Chinese (especially in Mandarin), has not made a clear statement about these relationships. Indeed, he weakens his thesis that Chinese is a polysyllabic and polymorphemic language by saying:

> The nearest equivalent to *tzyh* [tzu4] as a linguistic unit is the morpheme, which is usually defined as the smallest unit in the language which has a meaning. This is the sense in which Chinese has been called, and to a large extent is, a monosyllabic language - a language in which every syllable has a meaning. This so-called 'monosyllabic myth' is in fact one of the truest myths in Chinese mythology. Exceptions in which it takes more than a syllable to represent a minimum meaningful unit are relatively few. (1968:138-9)

Further, Chao says (p.140) that 'cases of morphemes less than a syllable are even rarer than polysyllabic morphemes', naming two which he considers 'the only important ones', viz., [-r] (see Table III, no. 2) and the suffix *m*. Clearly, Chao has overlooked the common occurrence of elision in Chinese, a frequent occurrence in an ordinary conversational style of speech. For example, in normal speech, the phrase *i4 ko4 jen2* 'a man', can sound like ['igrən] and not like [i gərən]; in such a context, the [g] is less than a phonetic syllable, although a full character and a full morpheme. Just as in spoken English, we expect elision and vowel reduction,[10] so also, in Chinese, contractions of full syllables to segments and prosodies must be expected, keeping in mind that the written language must still retain all the characters (*tzu4*) of the contracted utterance.

Most recently, Lin (1972) has voiced the importance of entering lexical words in a dictionary, but again, the individual entries in his dictionary do not always reflect the author's intention of keeping character,

Table II

Character/ syllable/ bound or free root/ affix 字 *ji4*	Morpheme				
	Bound root	Affix	Word		
			Free root (simple stem)	Compound stem	Stem + affix
1 *faad3* 法	*faad3* 'law'			*faad3*loed4 'law' faan4*faad3* 'go against the law' *faad3*gun1 'magistrate'	
2 *haay5* or *haay2* 鞋					
2a *haay5*			*haay5* 'shoe'	*haay5*pey5 'shoe leather'	
2b *haay2*	*haay2* 'shoe'			to1*haay2* 'slipper' law5baan2*haay2* 'indoor slipper'	
3 *jay2* or *ji2* 子 or 仔					
3a *-jay2*		*-jay2* (diminutive affix)			maaw1*jay2* 'kitten' gaw2*jay2* 'puppy'
3b *jay2*			*jay2* 'son'	*jay2*noey2 'children, offspring'	
3c *ji2*	*ji2* 'seed (of a plant)'			gwaa1*ji2* 'melon seed' lin5*ji2* lotus seed'	
4 *noey2* or *noey3* 女					
4a *noey2*			*noey2* 'daughter'	jay2*noey2* 'children, offspring'	
4b *noey3*			*noey3* 'female'	*noey3*jay2 'girl'	

Table II (continued)

	Bound root	Affix	Free root	Compound stem	Stem + affix
5 *yaw5* or *yaw2* 油					
5a *yaw5*			*yaw5* 'oil'	how5*yaw5* 'oyster sauce'	
5b *yaw2*	*yaw2* 'oil'			baag4*yaw2* 'soya sauce' din4*yaw2* 'petrol'	
6 *hung5* or *hung2* 紅					
6a *hung5*			*hung5* 'red'	*hung5hung5* 'reddish' *hung5*nay5 'laterite'	
6b *hung5*			*hung5* 'auspicious'	*hung5*yad4ji2 "auspicious date'	
6c *hung5*			*hung5* (edible blood of birds and animals)	jue1*hung5* 'pig's blood' gay1*hung5* 'chicken's blood'	
6d *hung2*	*hung2* (a shade of red)			jow2*hung2* (colour of dates) fan2*hung2* 'pink'	
7 *-did1* 的 or 啲		*-did1* 'a little' (adjec. affix)			hung5*did1* 'a little redder'
8 *-dey4* or *-dey2* 咗					
8a *-dey4*		*-dey4* (pronom. affix)			ngo3*dey4* 'we,us' koey3*dey4* 'they,them' yan5*dey4* 'oneself'
8b *-dey2*		*-dey2* 'a little' (adjec. affix			hung5hung5*dey2* 'a little red' ceng1ceng1*dey2* 'a little green'

Table II (continued)

	Bound root	Affix	Free root	Compound stem	Stem + affix
9 *-ga* 嘅		*-ga* (relative affix)[12]			ngo3*ga* 'that which is mine' din4ce1*ga* 'that which is the car's'
10 *maay5* 埋		*-maay5* (completive aspect)			hab4*maay5* 'to have finished closing'
11a *-faan1* 翻		*-faan1* (reciprocal aspect)			hab4*faan1*maay5 'to have finished closing again'
11b *-faan1* 番			*faan1* 'return'	*faan1*ley5 'come back' *faan1*hog4 'go to school'	
12 *si1* 斯 + *man5* 文			*si1man5* 'elegant, polite'		
12a *man5*	*man5* 'poetry'			*man5*hog4 'literature'	
13 *yi5* 而 + *gaa1* 家			*yi5*gaa 'now'		
13a *gaa1* 家			*gaa1* 'family'	*gaa1*je1 'eldest sister' *gaa1*po2 'mother-in-law' dong1*gaa1* 'to manage a household'	
13b *-gaa1* 家		*-gaa1* (agentive affix)			fo1hog4*gaa1* 'scientist' jing3ji4*gaa1* 'economist' yue5yam1hog4*gaa1* 'linguist'

syllable, morpheme and word distinct from each other. For example, (Lin 1972: 1160), *er1* [*erh2*] is entered as a non-syllabic form, particle, noun and adjective respectively. Although the first two entries recognize the non-word status of *erh2*, the last two entries do not; yet the citations under these last two entries are unsplittable compounds and reveal that *erh2* is a bound form (and therefore not a word).[11]

Finally, although tone modification is a very important feature of Cantonese, Cantonese lexicographers do not record it, with the result that many near-homophones, albeit semantically linked, are wrongly identified as the same unit.

4.1. Relationships in Cantonese The following points should be emphasized:

1) Homophones and near-homophones are not always clearly accounted for in dictionaries of Cantonese.[13] Sometimes, one entry is given and all its 'meanings' are listed under the one entry, with no clear indication of whether the entry is for a root or for an affix. For instance, nos. 13a and 13b *gaa1* and *-gaa1* are given one entry as a free form (Cowles, p.333 under *ka*) and glossed: 'Household. Family. Profession. Professional men. Experts. *S*[surname].' In Table II, I have tried to clarify problems of this kind by indicating the dual or multiple role of a particular character in terms of syllable, lexical root and affix. I have treated *-gaa1* in no. 13 *yi5gaa1* 'now' purely as a syllable (in this case without meaning); *gaa1* is a free root (and word) meaning 'family' in no. 13a; it is a free root (but not a word) in the compound *gaa1je1* 'eldest sister' (no. 13a); finally, in no. 13b, it is an agentive suffix in *fo1hog4gaa1* 'scientist'.

At other times, only one entry and one meaning are given in Cowles to a form which should be differentiated in some way from its homophones or near-homophones. For example, nos. 6c and 6d in Table II, *hung5* (free root) 'edible blood of birds and animals' and *hung2* (bound root) 'a shade of red' are not directly accounted for by Cowles, who gives one entry *hūng* [*hung5*], with the following gloss: 'Red. Lucky. Ruddy. *S*[surname].' *hung5* (no. 6c) is not even given under the relevant compounds, but *hung2* (no. 6d) is accounted for indirectly under the entry (Cowles, p.155) *fan-hūng* [fan2hung5], with the gloss: 'Pink. Flesh or salmon color.' No indication of the tone modification, is, however, given, so I have retained the unmodified tone 5 in my romanization within square brackets.

Similarly, nos. 8a and 8b, *-dey4* and *-dey2* are given the same character 𠱁 ; no. 8b is not given a separate entry but is entered incidentally under the head syllable in the phrase (Cowles, p.464): 'kwaai-kwaài-teî-téng-wâ [gwaay1gwaay1*dey2* teng1waa4] "be (*a little*) good and obedient".' However, in the case of nos. 11a and 11b, Cowles recognizes two forms, one free, one bound, but uses the word 'return' in both sets of glosses and shows their related meanings in the similar characters.

2) In some cases in Cantonese, a character which has been borrowed partly or fully from Mandarin is used to represent totally different meanings in Cantonese because the lexicographer, who is influenced by his subconscious knowledge of the existence of the styles of Chinese represented

in Table I, columns II and III, is unable to dismiss the Mandarin-influenced meaning from his mind.

Thus, the corresponding entries for *-did1* (no. 7) in Cowles are very confused. They are entered as free forms *ti* and *tik*, which Cowles identifies (p.994):

> *Ti*. Some. A little. A few. Makes possessive, comparative forms; and also participles and adjectives. See also *Tik*. Note 22.
>
> *Tik*. Clear. Time. A little or a few. Same as *ti*. See note 22.

Note 22 (Cowles, p.xvii), however, only deals with the Mandarin-influenced possessive meaning and does not clarify the status of the entry *ti* as an adjectival affix:

> 'Ti' (1) is a colloquial form now practically obsolete, 'ti' (2) being usually seen. This is pronounced 'tik' and thus appears in the kwŏk-uĕ [i.e., Mandarin] the same sense as 'ti' is used colloquially. When reading out from the kwŏk-uĕ text, it is often pronounced colloquially as 'ti'.

3) Certain characters are found in Cantonese but not in Mandarin (e.g., no. 9 *-ga* 㗎). Due to sociolinguistic factors, Chinese scholars in Malaya tend to avoid writing them down, preferring the Mandarin equivalent whenever possible; in this particular case, the Mandarin equivalent would be *-de* (see Table III, no. 5), but pronounced *dig1* (see Table I, no. 2 and remarks on *ti* and *tik* above). Perhaps for similar reasons, construction no. 10, *-maay5* is identified by Cowles (p.564) by implication with the free form *maaĭ* [maay5], being given the following gloss: 'With. Besides. To conceal. Bury.' This identification of two different forms is apparent from the fact that, on p.591, Cowles gives two entries *mi-maaĭ-háu* 'to purse the lips' [mi1maay5 haw2] and *mi-maaĭ-ngaăn* 'to close the eyes' [mi1maay5 ngaan5], in which the character for *maaĭ* [maay5] is the same as the one for *maaĭ* given on p.564 and quoted above. Clearly, only the form which is cognate with the Mandarin bound form *nieh1* (Mathews, p.654) has been consciously entered as a form by Cowles; the aspectual affix recorded in Table II is regarded as being relatively unimportant, and is only accounted for incidentally.[14] Yet the importance of differentiating between these forms is clear.

4) In nos. 12 and 13, we have examples of words of two syllables, each of which I have treated as monomorphemic words. Although this is not an analysis which would gain universal support, I nevertheless propose it because if each syllable (and character) were given a lexical meaning, the total 'meanings' of each word would be nonsensical: *si1* is entered in Cowles (p.844) as a bound form *sz* and glossed as: 'This. There. Any.' The *man4* part of the word is homophonous and homographic with the character which means 'poetry', and this could be of historical importance, but the total 'meaning', synchronically, would not be 'elegant'. A more sensible analysis would be to regard these two units *si1* and *man4* as syllables without individual meanings when they occur in the word *si1man4* 'elegant', especially

Table III

Character/ syllable/ bound or free root/ affix	Morpheme				
	Bound root	Affix	Word		
			Free root (simple stem)	Compound stem	Stem + affix
字 *tzu4*					
1 謝 *hsieh4*	*hsieh4* 'to thank'			*hsieh4hsieh4* 'thank you'	
2 兒 *erh2*					
2a *erh2*	*erh2* 'male'			*erh2*ma3 'stallion'	
2b *erh2*	*erh2* '(male) offspring'			*erh2*nü3 'offspring'	
2c *-erh2*		*-erh2* (dim. affix)			nü3*erh2*[ˌ....nyr] 'girl'
2d *-erh2*		*-erh2* (pl. affix)			hua1*erh2*[ˌ....hwar] 'flowers'
2e *-erh2*		*-erh2* (r-colour)			che4*erh2*[ˌ....dʒər] 'this'
3 水 *shui*			*shui3* 'water'	*shui2*kuo3 'fruit'	
4 果 *kuo3*					
4a *kuo3*	*kuo3* 'if, supposing'			ju2*kuo3* 'if, supposing'	
4b *kuo3* 果 or 菓	*kuo3* 'fruit'			shui2*kuo3* 'fruit'	
5 的 *-de*		*-de* (relative affix)			shui2kuo3*de* or [ʃuegwəd] 'pertaining to fruit'
6 麻煩 *ma2fan2*			*ma2fan2* 'troublesome'		

when *si1* is also used to transliterate the first syllables of 'Sparta' and 'Spencer' (Cowles, pp.849 and 850).

Similarly, if no. 13, *yi5gaa1* 'now', were treated as comprising two morphemes, we would have to account for the lexical meanings of *yi5* 'yet'[15] and *gaa1* 'family', resulting in a nonsensical 'compound'. Thus, it would be better (as I proposed in the first remark on Table II) to treat *-gaa1* as a syllable without meaning in the word *yi5gaa1*, and similary, to treat *yi5-* as a syllable without meaning in the word *yi5gaa1*. (See discussion of no. 6, *ma2fan2*, in 4.2. and see 5.0)

4.2. Relationships in Mandarin In 4.0, a passing mention was made of the unsatisfactory way in which *erh2* (nos 2a-2e) is entered in Lin (1972). In Table III, I have recognized two bound roots (nos 2a and 2b) with the general meaning 'male' and '(male) offspring' respectively. The bound root 2a *erh2* 'male' enters into the formation of compounds like the following:

erh2 + *ma3* 'horse' = *erh2ma3* 'stallion'
erh2 + *tzu3* (diminutive suffix) = *erh2tzu3* 'son'
erh2 + *hua1* 'flower' = *erh2hua1* 'male flower'.

The bound root 2b *erh2* '(male) offspring' enters into the formation of compounds like the following:

erh2 + *nü3* 'female' = *erh2nü3* 'offspring (of both sexes)'
erh2 + *fu4* 'wife' = *erh2fu4* 'daughter-in-law'
erh2 + *sun1* 'grandchild' = *erh2sun1* 'grandchildren (of both sexes)'.

In addition, in Table III, I have recognized three affixes (2c, 2d and 2e): a diminutive affix and a plural affix (which are nominal affixes), and a suprafix denoting *r*-colour which is not restricted to any particular part of speech. These affixes are identified as one affix (diminutive affix) by Cheng and include examples 33b - 46c in Cheng (1973), Ch.3. Such an affix 'loses its syllabicity and tone completely and becomes part of the preceding syllable' (Cheng 1973:25), so that the resultant words of two characters are monosyllabic words.

The diminutive affix *-erh2* (2c) enters into the formation of words like the following:

nü3 'female' + *-erh2* = *nü3erh2* [nyr] 'girl'
hsiao3 'small' + *-erh2* = *hsiao3erh2* [siaur] 'child'
nan2 'male' + *-erh2* = *nan2erh2* [nanr] 'boy'

The plural affix *-erh2* (2d) is used when nouns are referred to in a 'collective or plural sense' (Lin, p.1160), and it occurs in words like the following:

hua1 'flower' + *-erh2* = *hua1erh2* [hwar] 'flowers'
yü2 'fish' + *-erh2* = *yü2erh2* [jyr] 'fish'
jo1 'meat' + *-erh2* = *jo1erh2* [rɔr] 'meat'.

The suprafix *-erh2* (2e) enters into the formation of words like the following:

che4 'this' + *-erh2* = *che4erh2* [dʒɚr] 'this'

wan2 'enjoy (oneself)' + *-erh2* = *wan2erh2* [wɚr] 'to play; enjoy oneself'.

In Table III, no. 3, the free root *shui3* 'water' is coterminous with character and syllable. It either occurs as a word 'water' or it enters into the construction of the compound stem *shui2kuo3* 'fruit' by combining with the *bound* root *kuo3*,[16] which has the general meaning 'fruit'.

In examples 4a and 4b, the character *kuo3* represents two bound roots: 4a has the general meaning 'if, supposing' while 4b has the general meaning 'fruit'. The first takes part in the compound stem *ju2kuo3* 'if, supposing', while the second takes part in the compound *shui2kuo3* 'fruit'. In addition, 4b has an alternative character 菓 (i.e., 果 with 'grass' radical) which is not used to represent 4a.

In example 5, the relative affix *-de* is coterminous with syllable and character, but in rapid speech, it may lose its syllabic status and be reduced to a 'd'.

In example 6, *ma2fan2* 'troublesome' is identified with one area of meaning and is treated as a monomorphemic (but disyllabic) word because the individual lexical meanings of the characters, 'hemp' and 'troublesome' respectively, do not justify recognizing two separate morphemes.[17] Chao (1968:142), however, justifies his analysis of recognizing two separate morphemes by interpreting the first character as 'snarled like hemp',[18] but this is in fact not the *lexical* meaning of the character by itself (Mathews, p.608; Lin, p.888). When it occurs in compounds as a morpheme meaning 'hemp', it does not have the meaning 'snarled', e.g. *ma2pu3* 'hemp fabrics, linen'. I propose, therefore, that we regard a character sometimes as a meaningless syllable (as in *ma2fan2* 'troublesome') and sometimes as a (meaningful) morpheme (as in 麻布 *ma2pu3* 'hempen fabrics, linen' and 麻子 *ma2tzu3* 'hemp seed'). (See discussion of nos. 12 and 13 in 4.1.; see 5.0.)

5.0 General problems of word analysis in Chinese

From the foregoing arguments, we can see that certain disyllabic words were once probably dimorphemic and polymorphemic, but because of semantic 'fossilization' in the history of the Chinese language, we do not nowadays universally identify the individual syllables of words with separate meanings. Often, the whole di- or polysyllabic word is identified with one area of meaning. The first syllable of Cantonese *yi5gaa1* 'now' and of Mandarin *ma2fan2* 'troublesome', and the second syllable of Cantonese

si1man5 'elegant' could be examples of this problem. Not taking the individual 'meanings' of such syllables into account in a synchronic study, is, as I said earlier, analogous to deciding that the word *sinecure* in modern English is monomorphemic in spite of the fact that historically, two Latin words were involved. The problem, therefore, is not peculiar to Chinese linguistics, but it is a problem which can be contained as long as we do not confuse diachronic and synchronic treatments of a language.

The second problem is, perhaps, more exclusively Chinese in nature, although it exists to a lesser extent in other languages. This problem is directly related to the writing system of the language and it is a problem which I have tried to face in my treatment of the first syllable in *si1man5* 'elegant' and the second syllable of *yi5gaa1* 'now' in Table II. This problem centres on the fact that certain polysyllabic words cannot have had any individual meanings for every syllable, but the writing system, not being alphabetic, makes it possible for each character to be forced into a meaning mould if necessary although all that is required is a sound. For instance, as we saw in section 4.1., example 13 in Table II, *yi5gaa1* 'now' is written with the same character as the literary form of 'yet' 而 and 'family' 家 respectively. If we insisted that the 'real meaning' of *yi5gaa1* was 'yet family' instead of 'now' we would give a false picture of the language. We might get the same effect in English if we insisted that the word *sine* in trignometry was the same morpheme as S-I-N-E in the word *sinecure*, an analysis which cannot even be justified historically and one which would obscure the differences in pronunciation in the two sequences S-I-N-E.

A third problem of word analysis which is also related to the writing system of Chinese, a problem fully described in Chao (1946), is that of the 'loan character' and it illustrates incidentally the one-many and many-one relationship between the character and the word. In the history of Chinese writing, whenever an abstract idea was difficult to indicate by a pictograph, a character for a homonymous word with a picturable meaning was 'borrowed'. At a later stage, it became convenient to differentiate between two such characters by adding an additional element to reflect graphically their difference in meaning. For instance, Chao points out that the character *hui3* 毀 was once used to denote the meanings 'destroy' and 'destroy by fire', but later on, scholars added a fire radical 火 to the character in order to show the difference between these two meanings, with the following result:

毀 'destroy'

燬 'destroy by fire'

This kind of practice worsened when schoolmasters exercised their pedantry variously in order to overdifferentiate characters and make them appear more logical (see, for instance, Forrest 1958:38). This often resulted in the creation of larger and more complicated characters which became more and more difficult to memorize because of the increased number of strokes per character, a state of affairs which schoolmasters felt

provided excellent training for the young. In fact, this practice was in a sense a deviation from ancient scholarship, for in calligraphy, that very important part of Chinese art and scholarship, simple characters with only one or two strokes are the most difficult to write. If successfully written (i.e., with the brush), such characters are the most satisfying aesthetically, producing in the scholar/artist a mystical feeling of oneness with the character.[19]

5.1. Problems of mistranslation Another problem of word analysis in Chinese which justifies mention and some detailed discussion is the false view of Chinese which has been perpetuated by certain Western writers (notably Ernest Fenollosa and Ezra Pound), which view is criticized and dealt with at some length in Dembo (1963), Chapter 3. Briefly, this problem centres on the fact that a Chinese character may be made up of many *radicals*, each of which could occur by itself as a character in its own right in its basic form, when it would be a character made up of only one radical. A radical which is present in a character may indicate the general meaning area of the character; it may, by itself or in combination with one or more other radicals, indicate the general phonetic quality of the character (traditionally called the 'phonetic'), or it may be completely redundant. What Pound, under the influence of Fenollosa, did, was to recognize at times radicals and/or phonetic components within characters as words, ascribing 'meanings' to them, often embroidering the composite meanings thus produced with further meanings as fancy led him.

Perhaps Pound, who is not a serious scholar of Chinese, is not entirely to be blamed for his mistakes, for Karlgren (1923:4) says:

> The point on which the Chinese philologists as well as their western followers have entirely failed is the phonetic side of the study of the Chinese script; and yet it is obvious that just this side plays an extremely important part, as nine tenths of all Chinese characters consist of one 'signific' and one 'phonetic'. The Chinese scholars however have never treated these problems in a serious way, and their western followers have laboured under the difficulty that they have had to base their studies on modern dialects some 2000 years younger than Archaic Chinese [i.e., before 6th century A.D.].

Two examples of misanalysis and misunderstanding of the character by Pound appear in his translation of Confucian Ode 42 (quoted in Dembo 1963:27). They are the phrases *lady of azure thought* and *lady of silken word*, which are, respectively, the non-translations of the familiar word *ching4nü3* 'modest/chaste girl' and of the aphorism *ching4nü3 ch'i2lüan3* 'how beautiful is a modest/chaste girl!'. Unless mistranslations of this kind are exposed, Western literary and linguistic scholarship can be easily misled into false interpretations of Chinese thought and of the Chinese language.

In the first example of mistranslation, *lady of azure thought*, Pound was obviously playing around with the left-hand radical of the character *ching4* (in the compound *ching4nü3* 'chaste/modest girl'), at the same time bringing in some of the characters of which this radical forms

a part:

1 青 *ch'ing1*. This is Radical 174, which can occur as a character in its own right (and in an unchanged form) to mean (Mathews, p.164) 'the colour of nature; green, blue, black. A drab neutral tint'. It is its own phonetic component (Soothill no. 82)[20] and it is a near-homophone of *ching4* (no. 2) below.

2 靜 *ching4*. This is the character used in Ode 42 and it consists of (a) Radical 174 *ch'ing1* 青 on the left as the semantic component and (b) the phonetic component *ching4* 爭 (Soothill no. 413) on the right. The character 靜 can occur as a word which usually means 'quiet', but in the context of the Ode, it is a bound morpheme in the noun *ching4nü3* 'chaste/modest girl'.

3 清 *ch'ing1*. This character is made up of (a) Radical 174 *ch'ing1* 青 on the right as the phonetic component (Soothill no. 82) and (b) Radical 85 on the left as the semantic component. Radical 85, in its independent form *shui3* 水, occurs as a character in its own right and means 'water' (see Table III, no. 3). The character 清, a true homophone of (1) and a near-homophone of (2) above, means (Mathews, p.166) 'clear, pure, lucid'. It occurs as a word which is used generally, though not exclusively, to describe liquids.

4 晴 *ch'ing2*. This character is made up of (a) Radical 174 *ch'ing1* 青 on the right as the phonetic component (Soothill no. 82) and (b) Radical 72 *jih4* 日 on the left as the semantic component. Radical 72 occurs in an unchanged form as a character in its own right to mean 'sun, day'. The character 晴, a near-homophone of (1), (2) and (3) above, means 'clear, fair' and is generally used to describe the weather.

5 精 *ching1*. This character is made up of (a) Radical 174 *ch'ing1* 青 on the right as the phonetic component (Soothill no. 82) and (b) Radical 119 *mi3* 米 on the left as the semantic component. Radical 119 occurs in an unchanged form to mean 'rice'. The character 精, a near-homophone of (1), (2), (3) and (4) above, means (Mathews, p.161) 'the essence of; the essential' and is often used to describe abstract thought.

Clearly, the 'azure' of *azure thought* could have been suggested by the occurrence of Radical 174 in (2) (as Radical and semantic component)

and in (3) and (4) (as phonetic component), while 'thought' could have been suggested by the occurrence of Radical 174 in (5) (as phonetic component). Pound either ignored (or was ignorant of) the use of Radical 174 as a phonetic component. The arbitrariness of Pound's treatment of Chinese can be seen in the fact that the 'meanings' of the other radicals, viz., 'water', 'sun' and 'rice', play no part in his phrase *lady of azure thought*.

In the second mistranslation, *lady of silken word*, Pound was seeing too much in the fortuitous presence of individual radicals in the character *lüan3* 孌 'beautiful', a character which occurs in a stock phrase *ching4nü3 ch'i2lüan3* 'how beautiful is that chaste/modest girl!' [given on p.608 of Mathews in characters]. This character *lüan3* 'beautiful', is discussed below, together with a near-homophonous character *luan1* 'to tie together', to which it is graphically and phonetically linked.

1 孌 *lüan3* 'beautiful' consists of the character 䜌 (see (2) below) as a top phonetic element (Soothill no. 888) and Radical 38 *nü3* 女 as a bottom semantic element. Radical 38 can occur in an unchanged form as a character in its own right to mean 'female', a form which is cognate with Cantonese *noey3* (see Table II, no. 4). The character 孌 is listed under Radical 38 in dictionaries. Radical 38 is obviously one element from which Pound derives his 'lady' in *lady of silken word*, the other element being the *nü3* in *ching4nü3* 'chaste girl' mentioned earlier.

2 䜌 *luan1* 'to tie together' consists of two radicals: Radical 120 appears twice, once on either side of the character, and Radical 149 occurs between the two instances of Radical 120. *luan1* is listed in dictionaries under Radical 149 and acts as its own phonetic component (Soothill no. 888). Radical 120 and Radical 149 can occur as full characters on their own; when Radical 120 occurs as a character *ssŭ1* 糸, it changes its shape slightly, and can represent a bound morpheme, with the general meaning 'silk'; when Radical 149 occurs as a character in its own right (*yen2*) 言, it retains the same shape as the radical, sometimes representing a morpheme and sometimes representing a word; it then has the general meaning 'word' or 'speech'.

From the discussion of *lüan3* and *luan1* above we can get some idea of how Pound arrived at his phrase *lady of silken word*. However, we must remember that the incidental presence of Radicals 120 糸 and 149 言 in the character *lüan3* 'beautiful' is not significant from the point of view of translation because they are not directly concerned with the character *lüan3* 'beautiful'. They merely happen to be components of *luan1* 'to tie together', which, in turn, is the phonetic element of *lüan3*; thus *lüan3* is a combination of the whole of *luan1* as a phonetic element (not of radicals within *luan1*) and Radical 38 as a semantic element.

An analogy to Pound's treatment of Chinese would be an analysis of English in which we proposed a logical connection between the meanings of the words *shoe* and *hoe* by arguing that the sequence *-hoe* appeared in *shoe*. Further, we could argue that *shoe* and *hoe* must be deeply connected in meaning because a shoe touches the surface of the earth while a hoe digs the surface of the earth, these two functions (walking and digging) thereby illustrating the yeomanlike nature of the English man-in-the-street. This untenable view of English is not so absurd when we place it beside Fenollosa's interpretation[21] of the ordinary everyday Mandarin sentence:

人 見 馬

jen2 chien1 ma3 '(a) man sees (a) horse'.

> First stands the man on his two legs. Second, his eye moves through space: a bold figure represented by running legs under an eye, a modified picture of running legs, but unforgettable once you have seen it. Third stands the horse on his four legs.
>
> The thought-picture is not only called up by these signs as well as by words, but far more vividly and concretely. Legs belong to all three characters: they are *alive*. The group holds something of the quality of a continuous moving picture.

Here, Fenollosa appears to be partly looking at the pictorial origins of Chinese writing and partly looking for units of meaning which are smaller than the radical, as the three characters *jen2*, *chien1* and *ma3* are also three distinct radicals (nos. 9, 147 and 187 respectively). The top part of the character *chien1* (Radical no. 147) is the same as the character *mu4* 目 'eye' (Radical no. 109), which presumably explains why Fenollosa refers to an eye.

This misleading view of translation is by no means an outdated problem, because as recently as 1972, Barnstone made the same false assumption in his translation of Mao Tse-tung's poems when he said (Barnstone, 1972:141):

> The present versions may be called a close translation. My colleague Ko Ching-po and I have studied each character and then sought an English equivalent. Nothing has been omitted.

Barnstone, in his confusion between 'character' (which he calls 'multiple-meaning words' on p.141), 'word' and 'radical', misleadingly translates the character *ming2* 明 (which, in the context, means 'understand') into a nonsensical word 'sun-moon'. Mao's original line (Barnstone, p.80), 'Ch'ang2 yieh1 nan2 *ming2* ch'ih4 hsien4 t'ien1' '(In the) long night (it is) difficult to *understand* [i.e., visualize] a sky (as a) red province'[22] becomes Barnstone's 'Night is long. And slowly comes the crimson sun-moon dawn'. Barnstone attempts to justify this on p.129:

> This single character, one of several signifying brightness, contains the character for 'sun' next to the character for 'moon'.

However, from the foregoing argument, it is clear that for the word 'character' in Barnstone's explanation, we should read 'radical' and that we should not expect radicals to function as words in a sentence without great distortion of meaning.

6.0 Word and phrase

Finally, let us look at the problem of deciding whether a certain construction is an endocentric noun compound or a coordinative noun phrase. The Cantonese noun construction *hag1yan5* 'Negro', composed of a free adjective root *hag1* 'black' and of a free noun root *yan5* 'person', is a case in point. In my analysis, *hag1yan5* is regarded as a word, not a phrase, by appealing to the uninterruptability of the two components in the construction. It is true that *hag1* and *yan5* also occur independently of each other as adjective word and noun word respectively, and that potentially, *hag1* and *yan5* could occur as a noun phrase of two words, *hag1 yan5* 'a man in black; a man with a black face (or skin)'. Such a phrase could occur in a description of a play in which different people have their skins painted black, green and red, but in which all the actors are Chinese, no Negroes being present. If, furthermore, both the *ceng1 yan5* (the green man) and the *hag1yan5* (the black man) were fat, it would be possible to distinguish between them by interrupting the adjective and the noun words in the noun phrase with another adjective word *fey5* 'fat', resulting in the phrases *hag1 fey5 yan5* 'the black fat man' and *ceng1 fey5 yan5* 'the green fat man'. In the case of the noun compound, *hag1yan5* 'Negro', however, in order to say 'the fat Negro', one would have to put the word *fey5* 'fat' before the whole noun compound *hag1yan5* to form a modified noun phrase.[23]

This problem - the confusion between the word and the phrase - is made more complex in Cantonese because a large number of noun compounds are made up of two free noun roots, e.g., *din4ce1* 'motor car', *saam1doy2* 'shirt pocket' (see 3.2.), while one kind of coordinative noun phrase is formed by the mere juxtaposition of two semantically linked nouns.[24] Examples of such coordinative noun phrases are listed below:

1 saan1 soey2 'mountains (and) lakes; landscape'
2 toy2 dang3 'tables (and) chairs'
3 lung5 fung2 'the dragon (and) the phoenix; the groom (and) bride'
4 maaw1 gaw2 'cats (and) dogs'
5 fung1 yue3 'wind (and) rain; the elements'
6 haay5 mad4 'shoes (and) stockings'
7 naam5 noey3 'male (and) female; men (and) women'

The order of these nouns cannot be reversed in a phrase construction, although endocentric words could be coined, using the two roots in each case in reverse order, resulting in the potential forms:

1a soey2saan1 'a mountain made of water'[25]
2a dang3toy2 'a table shaped like a chair'
3a fung2lung5 'a dragon with the qualities of a phoenix'
4a gaw2maaw1 'a cat with the qualities of a dog'
5a yue3fung1 'a wind which presages rain'
6a mad4haay5 'shoes which also do duty as stockings'
7a noey3naam5 'a man who is not fully male'

A noun phrase of the type exemplified by nos. 1-7 above is capable of reduplication at phrase level to form a larger coordinative noun phrase, as in the following:

8 saan1 saan1 soey2 soey2 'lots of mountains and lakes'
9 toy2 toy2 dang3 dang3 'lots of tables and chairs'
10 lung5 lung5 fung2 fung2 'lots of dragons and phoenixes'
11 maaw1 maaw1 gaw2 gaw2 'lots of cats and dogs'
12 fung1 fung1 yue3 yue3 'wind and rain; all the elements'
13 haay5 haay5 mad4 mad4 'lots of shoes and stockings'
14 naam5 naam5 noey3 noey3 'everybody of both sexes'

Further, a noun phrase of the *saan1 soey2* type can be expanded into another kind of coordinative noun phrase in which a conjunction *tung5maay5* 'and' occurs:

15 saan1 tung5maay5 soey2 'mountains and lakes'
16 toy2 tung5maay5 dang3 'tables and chairs'
17 lung5 tung5maay5 fung2 'dragons and phoenixes'
18 maaw1 tung5maay5 gaw2 'cats and dogs'
19 fung1 tung5maay5 yue3 'wind and rain'
20 haay5 tung5maay5 mad4 'shoes and stockings'
21 naam5 tung5maay5 noey3 'men and women'

The sort of noun phrase exemplified by nos. 1-7 above may be initially confused with unsplittable endocentric noun compounds like *saam1doy2* 'shirt pocket' and *din4ce1* 'motor car' (see section 3.2). The two free roots (which are not semantically linked) in each compound can, of course, occur as words in sentences like:

saam1 yaw3 liong3 go3 *doy2* 'the *shirt* has two *pockets*', and
ce1 yung5 *din4* 'the *car* uses *electricity*',

but this is not the same as the occurrence of two roots within a compound word. Unlike noun phrases of the *saan1 soey2* 'mountains and lakes' type, noun compounds of the *saam1doy2* 'shirt pocket' type cannot undergo reduplication to form phrases of the kind exemplified by nos. 8-14 above; thus the following are not permitted forms:

*saam1 saam1 doy2 doy2
*din4 din4 ce1 ce1.

Only semantically linked nouns like *saan1* and *soey2* can participate in the two kinds of coordinative noun phrase illustrated by examples 1-7 and 8-14 above, the permitted order in which the components occur depending on

stylistic and other considerations which are outside the scope of this paper.

FOOTNOTES

* I wish to thank most warmly Dr. T. Bynon for the invaluable help which she has given me over the last two years, from the early preparation of this paper to its final completion. Without her constant encouragement and generous help, I would have been unable to finish writing it. I am also very grateful to Mr. J.R. Williamson for his kind help in reading, criticizing and correcting my near-final draft. I am further indebted to my former teacher, Mrs. N. Waterson, for her kindness in reading and criticizing an early version of this paper and for her continued exhortations not to abandon the work. My thanks also to Mr. J.S.-P. Lo for his beautiful calligraphy.

1 See Cowles (1965), p.xvii, notes 4, 12, 14, 16, 17, 18, 20, 21 and 22 for attempts to guide the student of Cantonese in reading, usage and pronunciation with this very problem in mind, although it is not explicitly referred to as a problem. Note 17, for example, says: 'In a great many expressions "loī" [loy5] may be substituted for "laī" (1) [lay5], "laī" being more colloquial, while "loī" is more on the bookish side, or found in more educated speech. However, there are some colloquial expressions where only "loi" can be used.'

2 A broad phonemic notation is used for these examples. The tone systems are explained in section 1.2. In the main part of this paper, which is not concerned with pronunciation, romanization is used instead of phonemic notation.

3 Although columns IV and VI appear to replicate each other, sometimes there is a spoken form which does not have a corresponding written form.

4 Robins (1964:333) appears to think that bound morphemes are rare in Chinese, but this is belied by the innumerable bound form entries in dictionaries. Cowles refers to them as bound forms and marks them with a plus sign although he does not always make a distinction between free and bound forms: e.g. Cowles (1965:994) *ti* [-dig1], an affix, is recorded as a free form.

5 For instance, *-ga* in the Cantonese relative phrase and *-m5-* in the Cantonese interrogative phrase. *-ga* and *-m5-* correspond to *-de* and *pu4-* respectively in Mandarin.

6 The two instances of *daay4-* are different prefixes and are written differently in Cantonese (see Table I, column V). The first mentioned *daay4-* is written with the character 大 while the second *daay4-* is written with the character 第 .

7 Lyons (1968:234) says of *longlegs* and *butterfingers*: 'The forms in question are not constructions at all; they are to be entered in the lexicon as unanalysable wholes . . . The distribution of

longlegs, unlike the distribution of *long legs* (in *he has long legs*) cannot be accounted for by any productive formational rules of modern English.'

8 In 3.5., *si1man5* is the only example of a monomorphemic word. The others, although undergoing the same reduplication, are words consisting of two roots because each part of the word is identifiable in other compounds and has similar semantic properties in the other compounds.

9 See previous fn.

10 For example, possible pronunciations of *I'm off* are [m ˈɔf] and [aIm ˈɔf], but not *[aI æm ɔf]. When *I'm off* appears in print as three orthographic words *I am off* (for instance, in a school reader), we still teach children to use the elided forms.

11 See Table III, no. 2.

12 This affix, like the Mandarin *-de* (Table III, no. 5), is glossed as a relative affix rather than a possessive affix because it is a level-skipping affix (typically, a phrase affix) which occurs at all levels of the grammatical hierarchy, turning whatever construction to which it is affixed into a relative construction. (See Killingley 1972:9.7.-9.7.2. and 11.1.)

13 As Cowles (1965) is the most comprehensive Cantonese dictionary available at the moment, I will refer Cantonese examples solely to it in this section.

14 I have not included a gloss 'bury, conceal' for the character *maay5* in Table II because in Malayan Cantonese the word for 'bury' is not *maay5* but *may5*. Note also that Cowles is inconsistent in recognizing a bound form corresponding to my no. 11a *-faan1* (reciprocal aspect) but in not recognizing a bound form corresponding to my no. 10 *-maay5*, although the two are parallel linguistic units, both being aspectual affixes.

15 This is a literary form not used in colloquial Cantonese. Again, the meaning of *yi5* could be of historical importance.

16 The cognate word *go2* 'fruit' in Cantonese is a free root but the Mandarin form *kuo3* is a bound root, occurring in compounds like *kuo2p'in3* 'fruits', *kuo2yüan3* 'orchard', *chung1kuo3p'i2* 'fleshy part of fruits' and *shih2kuo3* 'fruits in season'.

17 *fan2* also occurs as a bound morph with the meaning 'trouble' in productive compounds like *fan2hsin1* 'troublesome' and *fan2ch'i4* 'unhappy'. One solution is to treat *fan2* as a bound allomorph of the free morpheme (and word) *ma2fan2* when it occurs in such compounds.

18 This could be a consideration in a diachronic analysis of the word.

19 For a brief introduction to calligraphy and some account of the mystical experiences of the calligraphist, see Driscoll and Kenji Toda (1964).

20 'Soothill no.' refers to the phonetic index in Fenn (1940).

21 See Fenollosa (1936:8-9). In mitigation, as Fenollosa's essay was edited by Pound and published posthumously, Fenollosa cannot be blamed entirely for releasing the essay itself as a piece of scholarship.

22 The meaning of this line is: 'before the Revolution, it is difficult to understand (or see) the glory of China'. The word for 'sky' (*t'ien1*) also means 'day' and *ch'ih4 hsien4* 'red province' is obviously meant to make the reader think of an ancient name for China: *ch'ih4hsien4shen2chou1* 'the red land of the gods'. I have spelt *ch'ih4hsien4* here as part of a word but it is probably used in the poem as a phrase and as a pun (i.e. red province/dawn/the new China).

23 This is analogous to deciding that *blackbird* is a word but that *black* and *bird* can occur as words in a near-homophonous phrase. The word *blackbird* is uninterruptable, but the phrase *black bird* can be interrupted, e.g., *a black greedy bird; a black cruel bird.*

24 It would be interesting to examine the semantic relationships between such nouns and their classifiers. For a study of classifiers, see Coyaud (1973).

25 Only one possible meaning is suggested in each case. *soey2saan1*, for example, could also mean 'a mountain which gives water'.

REFERENCES

Barnstone, W. (tr.) (1972) *The poems of Mao Tse-Tung*. London, Barrie & Jenkins.

Bloomfield, L. (1933) *Language*. London, Allen & Unwin.

Chang, Hsüan (1968) *The etymologies of 3000 Chinese characters in common usage*. Hong Kong University Press. [Records the earliest forms of characters found on bronze and stone (c. 2000 B.C.) to the present day. In Chinese, but with a romanized index at the back].

Chao, Y.R. (1946) 'The logical structure of Chinese words', *Language* 22:4-13.

(1968) *A grammar of spoken Chinese*. Los Angeles.

Cheng, Chin-chuan (1973) *A synchronic phonology of Mandarin Chinese*. The Hague.

Cowles, R.T. (1965) *The Cantonese speaker's dictionary*. Hong Kong University Press.

Coyaud, M. (1973) *Classification nominale en chinois: les particules numérales*. The Hague.

Dembo, L.S. (1963) *The Confucian Odes of Ezra Pound: a critical appraisal*. London.

Dictionary of spoken Chinese. Institute of Far Eastern languages, Yale University Press, 1966.

Dow, Francis D.M. (1972) *The analysis of trochaic words in Chinese*. Edinburgh.

Driscoll, L. & Kenji Toda (1964) *Chinese calligraphy*. New York.
Fenn, C.H. (1940) *The five thousand dictionary*. Peking (1940 5th edition).
Fenollosa, E. (1936) *The Chinese written character as a medium for poetry* (ed. by Ezra Pound), San Francisco.
Forrest, R.A.D. (1958) *The Chinese language*. London.
Halliday, M.A.K. (1956) 'Grammatical categories in Modern Chinese', *Transactions of the Philological Society*, 1956:177-224.
Huang, Parker Po-fei (1970) *Cantonese dictionary*. Yale University Press.
Juilland, A. & A. Roceric (1972) *The linguistic concept of word*. The Hague. [A bibliography of material on the word in many languages.]
Kao, D.L. (1971) *Structure of the syllable in Cantonese*. The Hague.
Karlgren, B. (1920) 'Le proto-chinois, langue flexionnelle', *Journal asiatique* 15:205-32.
(1923) *Analytic dictionary of Chinese and Sino-Japanese*. Paris.
Killingley, S-Y. (1972) *The grammatical hierarchy of Malayan Cantonese*. Unpublished Ph.D. thesis, London.
Kramsky, J. (1969) *The word as a linguistic unit*. The Hague.
Kratochvil, P. (1968) *The Chinese language today*. London.
Lin, Yutang (1972) *Chinese-English dictionary of modern usage*. Chinese University of Hong Kong.
Longacre, R.E. (1964) *Grammar discovery procedures: a field manual*. The Hague.
Lyons, J. (1968) *Introduction to theoretical linguistics*. Cambridge.
Mathews, R.H. (1966) *Mathews' Chinese-English dictionary*. Revised American edition, Cambridge, Mass., 1966 (First published 1931).
Nida, E.A. (1949) *Morphology. The descriptive analysis of words*. Ann Arbor, (1949 - 2nd ed.)
O'Melia, T.A. (1961) *First Year Cantonese* parts 1-4. Hong Kong, Catholic Truth Society.
Pulgram, E. (1970) *Syllable, word, nexus, cursus*. The Hague.
Robins, R.H. (1964) *General linguistics: an introductory survey*. London.
Simon, H.F. (1958) 'Some remarks on the structure of the verb complex in Standard Chinese', *Bulletin of the School of Oriental and African Studies* 21:553-77.
Whitaker, K.P.K. (1955) 'A study on the modified tones in spoken Cantonese', *Asia Major* Vol. V, parts 1 and 2.